R E P O R T  R E S U M E S

ED 003 248 24

THE HIGH SPEED ELECTRONIC COMPUTER AS A RESEARCH AND OPERATIONS DEVICE IN SCHOOL LAW.

BY- ASHER, J. WILLIAM KURFEERST, MARVIN
PITTSBURGH UNIV., PA., SCHOOL OF EDUCATION

REPORT NUMBER CRP-1275 PUB DATE FEB 63
EDRS PRICE MF-$0.18 HC-$3.04 76P.

DESCRIPTORS- *COMPUTER PROGRAMS, *RESEARCH TOOLS, *LEGAL PROBLEMS, LAW SCHOOLS, LAWS, CODIFICATION, *INFORMATION RETRIEVAL, CONCORDANCE, PITTSBURGH, PENNSYLVANIA, KEYWORD-IN-CONTEXT (KWIC) PROGRAM

THE FEASIBILITY OF EMPLOYING ELECTRONIC COMPUTERS AS RESEARCH AND OPERATIONAL TOOLS IN THE AREA OF SCHOOL LAW WAS TESTED. TEXT MATERIAL WAS SELECTED, MARKED FOR TYPING BY AN ATTORNEY, AND TRANSFERRED TO MAGNETIC TAPE THROUGH THE INTERMEDIATE STEPS OF PUNCHED PAPER TAPE AND IBM PUNCH CARDS. A CONCORDANCE PROGRAM BROKE DOWN THE TEXT INTO ALPHABETIC WORD FREQUENCIES AND PINPOINTED EACH OCCURRENCE ACCORDING TO THE DOCUMENT IN WHICH IT APPEARED. THIS READY IDENTIFICATION ENABLED THE COMPUTER TO MATCH WORDS AND COMPARE DIRECTIONS GIVEN BY THE INVESTIGATOR. TWO BASIC PROGRAMS WERE UTILIZED DURING THIS PROJECT. THE FIRST WAS A STANDARD SEARCH WHICH INVOLVED THE COMBINING OF CONCEPTS AND HAVING THE COMPUTER MATCH AND LIST THOSE DOCUMENTS ACCORDING TO THE INSTRUCTIONS OF THE PROGRAM AND THE SEARCHER. THE SECOND PROGRAM, CALLED THE KEY WORD-IN-CONTEXT (KWIC) PROGRAM, ENABLED THE INVESTIGATOR TO IMMEDIATELY DETERMINE THE CONTEXT IN WHICH KEY WORDS WERE USED. A COMPARISON WAS THEN MADE BETWEEN MACHINE SEARCH AND A CONVENTIONAL HAND SEARCH DONE BY AN ATTORNEY. IT WAS CONCLUDED THAT RECODIFICATION OF LAWS IS FEASIBLE WITH COMPUTER ASSISTANCE, AND LEGAL INQUIRIES OF AN EDUCATIONAL NATURE CAN BE ECONOMICALLY SEARCHED BY MACHINE METHODS. (JK)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ED003248

# THE HIGH SPEED ELECTRONIC COMPUTER AS A RESEARCH AND OPERATIONS DEVICE IN SCHOOL LAW

J. WILLIAM ASHER, Ph.D. AND MARVIN KURFEERST, M.A.
SCHOOL OF EDUCATION
UNIVERSITY OF PITTSBURGH

The research reported herein was supported by a grant from the U. S. Office of Education - Department of Health, Education, and Welfare

COOPERATIVE RESEARCH PROJECT #1275

FEBRUARY 1963

# THE HIGH SPEED COMPUTER AS A RESEARCH AND OPERATIONS DEVICE IN SCHOOL LAW

Cooperative Research Project No. 1275

J. William Asher, Ph. D.

and

Marvin Kurfeerst, M. A.

School of Education

University of Pittsburgh

Pittsburgh, Pennsylvania

February, 1963

The research reported herein was supported through the Cooperative Research Program of the Office of Education, U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE.

* Computations were done by the University of Pittsburgh's Computation and Data Processing Center, which is partially supported by the National Science Foundation Grant No. G-11309.

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

3414
77

## ACKNOWLEDGEMENTS

The late Dr. Robert Brittell, Associate Dean of the School of Education, was originally one of the principal investigators and the project director of this research project. It was from his keen interest in the field of school law that this research grew. He quickly recognized the potential of his applied field for the use of the computer and the basic legal retrieval systems developed by Professor John F. Horty and his staff of the Graduate School of Public Health and by Professor William Kehl and his staff of the Computation and Data Processing Center. Dr. Brittell contributed greatly to the development, organization, and implementation of the research. It was his foresight and clear directives in planning that enabled the writers to complete this project. His presence has been sorely missed, not only by the investigators, but by his colleagues as well. We gratefully acknowledge his direction and leadership, and dedicate whatever value this report may have.

The basic research in legal information retrieval which preceded, paralleled, and continues is directed by Professor Horty, head of the University of Pittsburgh's Health Law Center. Without his complete cooperation, encouragement, and aid along with that of his staff: Seymour Schafer, Nathan Hershey, Eric Springer, Esquires; Mr. Thomas Walsh, and Mr. David Mitchell, this project could not have been done. In many respects this research was supplementary to his research. As his basic research in the field continues, undoubtedly it will open new research ventures for applied research in the field of school law.

Professor Kehl and Mr. Charles Bacon along with Mr. Joseph Naughton, all of the Computation and Data Processing Center gave every assistance to us in an area where few pathways have been developed. The computer analysis of language and words is an exceedingly complex problem. These men gave unsparingly of their minds and time to make feasible what had not been feasible before.

Our associates in the State Department of Public Instruction were most active in supplying problems and evaluating results of the output. This will become apparent as the report is read. Mr. John D. Killian III, Deputy Attorney General of the state of Pennsylvania and principal legal advisor to the Superintendent and State Department of Public Instruction, worked very closely with us and assisted us in disseminating our research in the state. Dr. Neal Musmanno, Deputy Superintendent of Public Instruction and Mr. Robert Broughton, Deputy Attorney General, also worked with us.

Other professors at Pitt who contributed were Drs. Samuel Francis, Richard Wynn, Herman Grose, and Richard DeRemer. Mr. Niles Anderson, School Solicitor for the Pittsburgh Public Schools

also consulted with us in connection with the use of these techniques for local school systems. The members of a school law class, Educational Administration 257, acted as "guinea pigs" for several of our research procedures.

In the United States Office of Education two staff members concerned with problems of school law visited with us and continued to confer with us from time to time. These men are Dr. James Gibbs and Mr. August W. Steinhilber.

Dean Paul H. Masoner of the School of Education proved beyond any doubt that he was willing to give every assistance to the research efforts of his faculty. Such Deans are rare and his assistance is appreciated.

Finally, could any research project ever be conducted, to say nothing of reported without the efforts of devoted secretaries? Ours were Mrs. Lois Churchill and Mrs. Estelle Goodykoontz.

J. William Asher
Marvin Kurfeerst

TABLE OF CONTENTS

List of Figures . . . . . iii

I. Introduction . . . . . 1

A. Previous Research . . . . . 4

II. Preparation of Materials . . . . . 6

III. Preparation and Initial Evaluation of Inquiries . . . 17

A. Source of Inquiries . . . . . 17

B. Framing of Inquiries . . . . . 19

C. Professional Evaluation of Inquiries . . . . . 23

IV. Results and Discussion . . . . . 28

A. Standard Searches . . . . . 34

1. E 9 . . . . . 34

2. E 10 . . . . . 35

3. E 12 . . . . . 36

4. E 14 . . . . . 38

5. 23 . . . . . 39

6. 75 . . . . . 41

7. 81 . . . . . 42

8. 83 . . . . . 43

9. 89 . . . . . 44

10. 179-180-181 . . . . . 46

11. 182 . . . . . 47

12. 183 . . . . . 49

13. 196 . . . . . 50

B. Keyword-in-context Searches . . . . . 51

1. KWIC 1 . . . . . 51

2. KWIC 2 . . . . . 51

3. KWIC 3 . . . . . 52

4. KWIC 4 . . . . . 53

5. KWIC 5 . . . . . 54

V. Potentials of Computer Analysis of Documents . . . . . 56

VI. Conclusions . . . . . 63

References . . . . . 67

## LIST OF FIGURES


Figure 1 Data Preparation Flow Chart . . . . . 8

Figure 2 Search # 81, Example of Framing . . . . . 21

Figure 3 Manual Search Index Entries (#81) . . . . . 24

Figure 4 Body of Manual Search (#81) . . . . . 25

Figure 5 List of Standard Searches . . . . . 29

Figure 6 List of KWIC Searches . . . . . 30

Figure 7 Standard Search Evaluations . . . . . 31

# CHAPTER I

## INTRODUCTION

This project was undertaken to test the feasibility of employing high speed electronic computers as a research and operational tool in the area of school law.

The major assumption underlying the need for such an investigation is that it is no longer simple and practical to continue using conventional procedures and materials to sort and find pertinent legal data. The justifications for such an assumption is based upon the following considerations.

However, initially it must be noted that the nature of legal material makes it subject to human error. Law itself is an expression of the human mind. Classification and index systems, however ingenious and helpful, are devised by humans and used by humans. They are subject to the limitations of human abilities, time, and initiative. Today the volume of legal material is increasing as never before. Judicial opinions alone number approximately 2,200,000 documents and are increasing at a rate of over 25,000 per year. Over 10,000 volumes contain these opinions. The day is past when even the country lawyer could wade leisurely through the many classifications of the law in the hope of locating all relevant material.

The first of these considerations is that the law, itself, is an appropriate field for data retrieval. The statutes, especially, are written in relatively clear, concise language. At least this is their

intent. Practically, this means that input and output can both be relatively short and that retrieval of legal information will be involved with fewer semantic difficulties. In addition, there exists in each of the fifty states and in combinations of them central depositories of similar information, organized almost identically. Only through exhaustive research and effort could this be accomplished in other areas. This pre-compilation organization of material is an especially significant point in the selection of legal data for use in experimental information retrieval programs.

A second consideration rests upon the nature of legal searches. No question can be answered until an exhaustive review of the data bearing upon the question has been completed. Whether it is a question regarding an individual or one that bears upon the future of us all, a similar searching procedure must be undertaken. The speed and degree of success depend upon the amounts of money and time available, the resourcefulness of the searcher, and the scope of the problem. Since all of these factors are involved, new avenues which will accelerate the answers to legal questions appear to be needed. It is felt that legal information retrieval through the use of the electronic computer offers an avenue for success because of the speed and ultimate completeness possible. For reasons mentioned earlier, the hand search presents too many opportunities for human error. Information retrieval should help to lessen and perhaps even remove these often costly experiences.

The third consideration of law as an appropriate area for information retrieval concerns the methods for indexing currently being used. We must return to the human limitations built into legal indexing systems.

All too often, a vital document is unidentified because of the indexer's human inability to adequately determine the correct category in which the document should be placed. Furthermore, legal indexes are limited in descriptive categories, and all too often documents are placed under headings which are, at best, ambiguous. The lawyer frequently must settle for less than a thorough search, although he has carefully and diligently pursued his task using all the standard resources. The time consumed, the built-in-biases of the reference systems, and the possibility of human error all mitigate the accuracy, speed, and success of the human effort.

The use of the computer as legal indexer, with its ready access to complete textual material, is well within the realm of consideration. Its potential uses in this respect include the functions of all present indexing systems and await only the ability of the programmer to develop more.

The final consideration might be characterized by the word "growth." With a system of information retrieval there appears to be unlimited horizons as to the absorption of new data. This includes not only additions of statutes and judicial opinions which are certain to increase, but also new "law-making" agencies such as state and federal commissions and regulatory bodies. It is anticipated that further technical developments in the next few years will permit us to incorporate far more material than the present computers are able to do. Therefore, the computer exemplifies a system which will not require completely new revisions and editions every time that new legal information is added. These assumptions, if true, are obviously equally appropriate to school law.

Previous Research

In 1955 an initial start was made at the University of Pittsburgh to develop a Data Processing and Computing Center using the most rudimentary type of equipment. By 1959, however, the concepts and practices of information retrieval had become an integral part of the Center. Working with an IBM 650 with unit tape storage, the University moved rapidly to an IBM 7070. There is now in operation a supplementary IBM 1401 which greatly increases the speed and lowers the cost of input and output of the data. (Already plans are made to obtain even faster and larger equipment.)

Early experimentation in legal searches with the electronic computer was done by Professor John F. Horty (4,5,6,7,8,9) and Professor William B. Kehl (13) at the University of Pittsburgh. This was done as part of a project in the Graduate School of Public Health designed to study and improve the health statutes of Pennsylvania. In March of 1960, the Health Law Center of the University agreed to give a demonstration of information retrieval for members of the American Bar Association at their August meeting in Washington, D. C. In conjunction with the American Bar Association and the National Institutes of Health, demonstrations of legal information retrieval were given over a nine day period. Because of space limitations, approximately thirty people per demonstration were permitted, yet a total of over 1,300 was accomodated during this demonstration including the United States Commissioner of Education and the Deputy Commissioner.

As a result of the enthusiasm generated by these demonstrations, a special committee of the American Bar Association was elevated to regular

status. This committee sponsored a second successful demonstration at the American Bar Association Convention in St. Louis during the summer of 1961, and they met in October of 1961 on the Pitt campus to make further plans.

The Council on Library Resources, Inc., in 1961, made a grant to the Health Law Center for the study of search techniques applicable to the entire statutory compilation of a particular state and the development of certain sophisticated search programs and techniques. The Ford Foundation also made a grant to the School of Law to support a detailed five-year study of the application and effect of information retrieval techniques upon the administration of justice and the legal process itself. Much of the research under this grant is being done by Professor Horty and his staff, several of whom hold faculty appointments in the School of Law as well as the Graduate School of Public Health.

Several papers have been published describing facets of the legal information retrieval research program at the University (1,6,7,8,13). One recent paper consists of a statistical evaluation of the retrieval system (14).

## CHAPTER II

## PREPARATION OF MATERIAL

The mathematical and data analysis procedures and capabilities of the computer are among its most widely used functions. However, computers can perform logical operations on verbal materials by comparing various symbols and indicating whether they are identical or not; then sorting them into piles of identical symbols and counting them. It can even "remember" where within the material the symbols came. Thus, the name "computer" could very well be "logicer."

Letters are symbols, and words are made of letters and spaces. The computer can compare these symbols when combined into words, count them, and remember where they occurred. Computers receive information from which to work from magnetized spots on magnetic tape, much like the tape on a home tape recorder except it is of better quality. These magnetic spots, or lack of spots, are detected and put into the logic and counting mechanisms of the computer. Here, the various spot combinations representing letters, numbers, punctuation and spaces can be "identified", compared with previous symbolic combinations, counted, and located. Large combinations of symbols, such as a sentence are recognized by the computer as a sentence whenever a period, question mark, etc. is followed by two blanks. Thus, a complex and rapid tool for analyzing the printed aspects of language is available. The number of times each word is used, the total number of words, or the total number of sentences, can be readily counted. With the ingenious use

of the computer a great diversity of quantitative and locative information can be gathered, sorted, and otherwise organized (13).

Since major costs and problems are associated with the preparation of materials for computer use, the procedures developed at the University of Pittsburgh will be summarized along with possible modifications that might be made because of newer equipment now available.

In the process of preparing documents for retrieval, the following steps are completed. Any competent typist can be employed for the initial phase. The documents are marked initially -- for this project an attorney did the marking to show the typist what is to be included; e.g., in legal material the title, citation, statute number, date, etc. Next, the typist -- in the current procedure using a machine known as a Flexowriter -- actually types the document on special sheets of typing paper which have carbon paper separating the sheets. The Flexowriter, which looks like a common, electric typewriter, differs in one aspect. As the typist is copying the document, the same document is being transposed into a code which is punched on a roll of paper tape mounted alongside. Errors which are noted on the page during typing can be corrected on the paper tape. This completed tape represents the document in its initial coded form.

The next step is to get the information from paper tape to magnetic tape. (Figure 1) At present, there is an intermediate step in this conversion; that of transferring the coded information from the paper tape into IBM punch cards, and then from the cards to the magnetic tape.

## DATA PREPARATION FLOW CHART

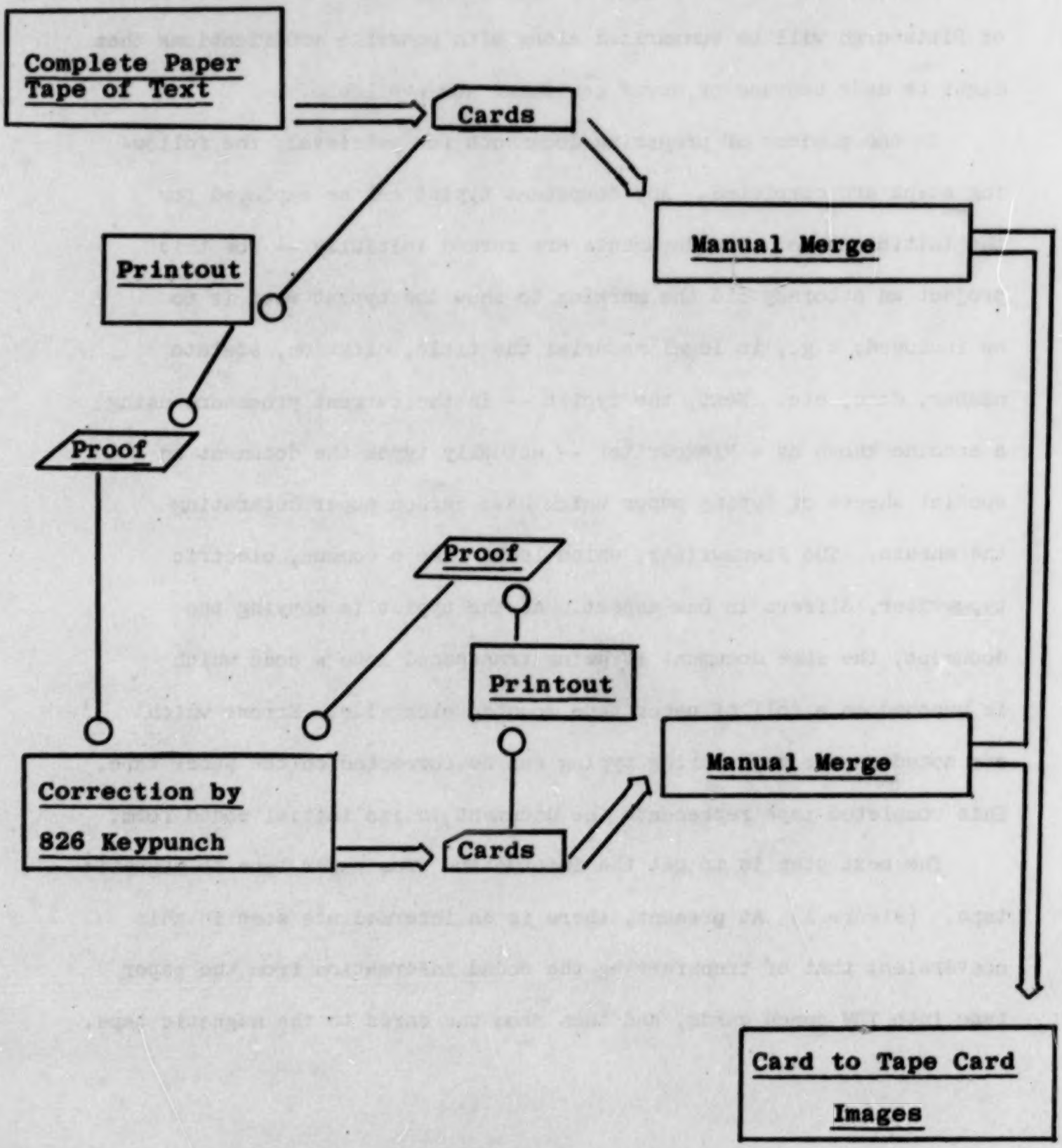


Figure 1

This step is somewhat expensive and time-consuming. The paper tape is mounted on a console of an IBM 047 machine, the appearance of which resembles a standard IBM keypunch machine. It has a typewriter on the front, a tray on the right for placing blank cards, and a tray for the completed cards on the left. As the paper tape is read, the IBM cards are punched at an average rate of 1000 cards per hour, with the characters from the paper tape being punched into the IBM cards -- the words at the top of the cards. This again facilitates any corrections, since any card having an error can be replaced immediately with an individually corrected card. Because the punched paper tapes are created expressly for conversion purposes, each Flexowriter line corresponds to one punched card. (In practice at the University it is limited to 75 columns with the last five columns, 76-80, reserved for machine processing codes.)

The completed cards are then placed in an IBM 407 tabulator-printer which reproduces a copy of the original document for proof-reading purposes. Correction editors are given a copy, they are proofread and marked, then new cards are punched. A second printout is obtained, proofread, and further errors corrected. Once the completed deck has been verified for accuracy, the cards are then loaded into the IBM 1401 to transfer the information from the punched cards to the magnetic tape. These magnetic tapes containing the reproduced documents in the form of magnetic "spots" on the tape are the Text Tapes.

All that has been done was to transfer the documents desired for inclusion from the printed page to magnetic tape. This seemingly simple process is in fact one of the major aspects of document retrieval and word analyses because it is time-consuming and relatively expensive.

The editing process on the original documents is not too difficult but must be done by rather able and therefore reasonably well paid people, usually professionals in the field from which the documents come. The typing of the documents is not difficult and can be done by personnel skilled only in operating a typewriter accurately and with some speed. However, the procedure can go no faster than a typist can type, although several girls can start at different places in a document.

The real cost, and one of the major problems of the process, lies in the extreme accuracy which must be achieved in the proof reading of the documents. The punched paper tapes, the original typed sheets, and the correction cards all are aimed at approaching absolute accuracy. This increases costs. The computer, while exceedingly fast, accurate, and moderate in cost, is exceedingly "dumb" at least by the human standards. If the symbols it receives spell "A-C-U-R-A-C-Y," it doesn't know that the two "C's" are redundant. Most people automatically will read the word as "ACCURACY." Instead the computer will frantically compare the symbols "A-C-U-R-A-C-Y" with all the word symbols previously received. When it finds no identical matches, it will announce its discovery by printing "ACURACY" on the word list along with the location in the documents and the number of times it found it, hopefully once. This human inaccuracy requires extra effort for all procedures which follow. For instance, if the use of the word "accuracy" is required, "acuracy" must be included in the list of search words.

The above procedure is one of the major costs of this type of language analysis. A rough estimate of changing the printed document into magnetic tape suitable for use by the computer can be put at about a

half cent per word exclusive of overhead costs. Once, however the tapes exist, they can be duplicated readily, and a system of exchanges could be developed so that documents need be reproduced once.

There are also indications that costs might be reduced. For example, the punched card stage might be skipped by converting the punched paper tape code directly into magnetic tape with the proof sheets being produced later from either tape. The corrections could be made by an error correction program for the computer. Also, the IBM 407 tabulator-printer, now used to print the information from the cards to sheets for proof reading, is rapidly becoming obsolete. The combination of the smaller IBM 1401 computer and newer high speed printers is faster and cheaper.

For the future there are two additional hopes. One, is when the documents are originally being set into type, punched paper tapes or for that matter magnetic tapes, will be used to concurrently store the information in the documents and thus make it readily available for use by a computer. Second, that equipment will become available to "read" printed documents directly and convert them into magnetic tapes. Such equipment can already read numbers, some even handwritten numbers (12, p. 5). The first hope is already with us to some extent in that some printed materials already are transmitted over wire service facilities in a form which could be readily converted to a form useable by a computer.

To this point, the procedure of transferring the printed page onto magnetic tape has been described. The final steps prior to the retreiving of information are briefly described below. A detailed description of the procedures has been published by Kehl, et al (13).

The cards can have text punched only in the first 75 columns. The

last five columns are reserved for special coding. In the system at the University of Pittsburgh the 77th column describes the type of material from which the documents come; citations, description, text, or annotation. Each of the four types in this case being a particular part of the statutes. Since these four aspects of the statutes are coded on the input with separate identities, each may be identified. Such a breakdown enables the machine to provide upon request; (1) printouts of serial numbers arbitarily assigned to the documents, (2) citations of statutes, sections, titles, etc., and (3) printouts of entire texts of documents.

The next major step required is the preparation of a Vocabulary Tape. The computer alphabetizes all the words on the text tapes into alphabetical order and in the process counts the frequency and location of each word's occurrence. This breakdown allows for even further pinpointing, namely, the location of each word as to sentence, paragraph, and placement within the sentence. Further, when each document was transferred to magnetic tape, not only were the original legislative numbers recorded, but the computer inserted a sequential number for each document, e.g., 1125, 1126, etc. These numbers are also listed for each word.

To save output time and space, certain so-called common words or numbers which have no major contribution to make to searching are omitted. At present, about 112 such words, generally articles, common prepositions, conjunctions, and numbers are omitted. This Vocabulary, or Concordance, Tape minus the common words then becomes the Basic Search List.

In summation then, the documents are transferred from paper tape to

cards to magnetic tape, which forms the Text Tape. Finally, an alphabetical Concordance Tape is created, which, in essence, becomes the Basic Search List.

An illustration of an actual problem will prove helpful. Remember that the Vocabulary Tape, or Basic Search List contains all the words of the original text according to location by title, document number, sentence, and sentence position, the first step in retrieval is to use those words which will elicit the greatest number of relevant documents. In order to cover all the words that might be used, including plurals and misspellings, a visual search is made of the printout of the Concordance. Each word is cited alphabetically, followed by a series of numbers giving its exact location. For example, the word "parochial" used in the search described below, might have the following series of numbers:

| | | |
|---|---|---|
| Parochial | 1106.001.5 | 1239.017.9 |
| | 1399.022.6 | .480.003.4 |

Reading the printout, "parochial" appears in document 1106, sentence 1, as word 5. It appears again in document 1239, sentence 17, as word 9, etc. Several variations of the same word are often found due to misspellings in the documents. For search purposes, these misspelled words must be included in the search, or the searcher runs the risk that certain documents will not be included. For example, "parochial" might appear as "parochiol," or "porochial."

Now for an example of a search. The question posed was, "What powers do the Superintendent of, the Department of, and the State Council of Education have over private and parochial schools?" First, the ques-

tion had to be rephrased for machine use. Instead of a sentence, groups of words covering all of the concepts involved had to be noted. For this question, four basic concepts were involved. First, schools which receive state assistance; second, the various types of religious and private schools; third, the various levels of education; and fourth, the three controlling agencies of education within the state.

The first concept involved the words "partly," "supported," and "commonwealth" with the word "partly" appearing beside "supported," and both words appearing in the same sentence with "commonwealth." Retaining this concept internally, the machine next searched for the second and third concepts, that of private and religious schools with their various levels and types. The documents containing such words as "parochial," "sectarian," "colleges," "private," etc., were also stored in the machine. The fourth concept, that of the Department of Public Instruction, the Superintendent of Public Instruction, and the State Council of Education, was framed so that in each phrase such as "superintendent, public, and instruction" are in the given sequence.

Holding these concepts temporarily, the storage unit of the computer was instructed to match in the same document the combinations of concepts programmed by the searcher. In other words, the concept of the controlling state agency had to appear in the same document with a merger of the other concepts involved. This merging of concepts is done by the computer in seconds. Therefore, the matching process, although appearing to be a cumbersome operation, is swiftly done.

Without technical details, the computer does the search in the following manner. The Vocabulary Tapes, or Basic Search List, are loaded onto

magnetic tape drives. While the loading is being done the computer program to be followed is loaded on another magnetic tape drive. (This program had previously been punched into cards with the list of words and the directions to be followed included before being transferred to magnetic tape.) While the tapes are being run and the machine is "digesting" the words and directions, a blank or "scratch" tape is loaded on another tape unit. When all the tapes have been searched, the computer will acknowledge its acceptance of all of the data from the cards and tapes by typing out a code of acknowledgement.

With all the data now present in the storage units of the computer, the machine begins to match the concepts with the documents as programmed. As each match is made, the information is transferred to the scratch tape. When the computer acknowledges that it has completed its searching and matching, "cite" and "print" commands to the computer are then executed. "Cite" involves printing the serial number of the selected documents, and "print" the title or full text. Depending on the purpose for which the search was undertaken and the "execute" commands, the computer will print the serial numbers of the documents in which all the required concepts have been met, the legislative numbers given to the statutes, the full titles of the statutes, or the entire test of the documents.

Originally the time required to search through the entire Vocabulary Tapes for the Pennsylvania Statutes, to prepare the search structure that results in a list of relevant documents, and to print out citations to these documents was between forty-five minutes and one hour of computer time on the IBM 7070.

It was through these procedures that the Education statutes of

Pennsylvania and the Attorney Generals opinions germane to Education were prepared for computer searching. In addition to the volume of the Pennsylvania statutes relating directly to education, some other miscellaneous educational statutes were also prepared. (It was originally planned that the Rules and Regulations of the Pennsylvania State Department of Public Instruction would also be put on tape. However, it became imperative that use be made of the tape library even though the rules and regulations of the state department were not yet ready for publication. Thus they could not be a part of the taped library.) Actually, since this research in school law supplemented a more general, ongoing research effort in law, advantage could be taken of the remaining Pennsylvania statutes which were also put on tape. This meant that all statutes which could even remotely influence education were available. Thus a realistic test of search procedures could be made.

## CHAPTER III

## PREPARATION AND INITIAL EVALUATION OF INQUIRIES

### Source of Inquiries

As envisioned in the proposal, it was felt that a major source of requests for searches would come from school solicitors. It was recognized that perhaps the greatest service could be to those agencies located in Harrisburg; namely, the Department of Public Instruction, the Office of the Attorney General, the Legislature, and the State Council of Education. As will be noted later, significant assistance was given to the first two, and great promise was evidenced for the third. On the local level, however, cooperation was extended by the School Solicitor of the Board of Public Education of Pittsburgh and with several local districts. Also, graduate students with special interests in school law, as well as the professional staff of the School, were provided service.

Despite a reasonably extensive publicity campaign in area papers and through the information channels in education, the local school districts, represented by their administrators and attorneys, contributed few useable questions for searches. This was not because of lack of interest. A number of questions were submitted but most dealt with the interpretation of a single law or with court opinions. The interpretations, of course, need to be made by lawyers or by the state department officials. Furthermore, this type of information was not yet a part of the materials accessable for computer searching.

It should be noted that the assistance rendered to these people was overshadowed by the contributions made by them to our project. By supply-

ing questions whose answers could be of value in a "practical" sense, much of our experimentation was made easier. For example, a professor was interested in knowing if two or more school districts could jointly purchase certain supplies and materials. The need for this answer and its accessability by computer was the test. In this instance, the appropriate answer was supplied which completely justified the original supposition. Several other such requests by the professional staff were similarly treated.

Due to the interest in the project, local school districts often phoned or wrote for information. In some instances, questions were submitted on initial inquiry, and in others, in subsequent correspondence. In return for providing them with experimental evaluations, some "feed-back" as to the degree of help was submitted.

At this point, it would be proper to acknowledge those people who met informally and provided the project with much help and stimulation. The primary group, comprised of those interested in the application of electronic techniques to retrieving school law, was organized by the late Dr. Robert W. Brittell, Associate Dean of the School of Education. In addition to their attendance at conferences held to acquaint them with the uses of the retrieval system, individual members of the group submitted questions which reflected problems in areas of school law with which they were acquainted. The work group included among others: Dr. C. Herman Grose, former school superintendent, Deputy Superintendent of the Pennsylvania Department of Public Instruction, President of California State College (Pa.), and presently head of the Office of Educational Placement; Mr. Niles Anderson, Solicitor for the Board of Public Education

of the City of Pittsburgh; Dr. Richard Wynn, Associate Dean of the School of Education and former teacher of Pennsylvania school law; Dr. Samuel Francis, former assistant supervising principal, and whose dissertation was in the area of school law, and is the instructor of courses in school law; Mr. Joseph Werlinich, doctoral candidate and former administrator in higher education; Dr. Stephen Storkel, who is now a principal in the Evanston (Ill.) Township High School; and Dr. Richard DeRemer, Associate Professor of Educational Administration, University of Pittsburgh.

The second group whose assistance proved invaluable, was the school law class itself (Educational Administration 257). Their patience and enthusiasm merits the individual listing not permitted in such a short space.

Framing of Inquiries

In those instances where an inquiry was received from someone within the School of Education, the person in charge of framing the inquiry attempted to determine and pinpoint the issue in question. By doing this, both the framer and the inquirer narrowed or broadened the inquiry, and suggested possible lists of words that could conceivably be used in the search. Furthermore, the various possibilities of grouping the words were explored so as to provide the optimum conditions for retrieving documents bearing upon the question. When the inquiry was received from an outside source, the framer of the inquiry was largely left to his own interpretations as to choice of words and/or grouping of words.

Therefore, the first step in framing an inquiry was to divide the questions into its component concepts; i.e., the different groups of people, types of action, objects involved, etc. Then, the framer listed

words under each concept which may be thought of as being synonyms, yet according to a strict definition, are not. Instead, these words are referred to as "searchonyms." An example of such searchonyms would be "questionnaire," "survey," and "experiment" under the broad category of research. In other words, the intent was to include words which would elicit all facets of the concept within the inquiry.

The next step was to check the frequency list of the Statutes in order to check the frequency of occurrence and, just as important, to note the various forms of misspelling which occurred as a result of legislative misprinting or incorrect key-punching. When all possible variations had been noted, the framer decided the relationship in which these concepts were likely to be found in the document.

The next step was to utilize a set of commands by which the various concepts could be linked. These commands had been devised by the programmers of the Computing Center as part of the overall search program. This is reported in an article by Kehl, Horty, Bacon, and Mitchell (13). and is briefly noted below.

> The four basic "relational operators" most commonly employed were:
> (1) A plus sign (+) meaning, "or," links any given group of searchonyms into a single merger or one list.
> (2) Capital "D" links any two concepts so that they must appear within the same document.
> (3) Capital "S" links any two concepts so that they must appear within the same sentence.
> (4) Capital "W" links any two given concepts so that they must appear within a specified range of words of each other.
>
> An illustration of the above is noted in Figure 2.

UNIVERSITY OF PITTSBURGH
INQUIRY #81
MACHINE SEARCH BY: Kurfeerst
QUESTION: Is there anything in the law that requires a board of school directors to offer a kindergarten or to offer a high school program?

| | | | |
|---|---|---|---|
| Kindergarten | + | Kindergarten | Print |
| Board | + | Board-s | |
| | + | Boards | 81A |
| Director | + | Director-s | |
| | + | Directorn | |
| | + | Directors | |
| | + | Directros | |
| | + | Diredctors | |
| | + | Direectors | 81B |
| Districts | + | Distrct | |
| | + | Distrcts | |
| | + | Distric | |
| | + | District | |
| | + | District-s | |
| | + | Districts | |
| | + | Dstrict | 81C |
| High | + | Secondary | |
| | + | Senior | 81D |
| Schols | + | School | |
| | + | Schools | |
| | + | Schoolship | |
| | + | Schoool | 81E |
| 81D | S | 81E | 81F |
| 81A | w+2+2 | 81E | |
| | S | 81B | |
| | + | Schoolboard | |
| | D | 81F | Print |
| 81E | S | 81C | |
| | + | School-district | |
| | + | School-districts | |
| | D | 81F | Print |
| 81D | w+1+2 | 81E | Print |

Figure 2

Another simple procedure for framing inquiries is the KWIC search. The KWIC, Keyword-in-Context Program, has proved to be extremely valuable in a variety of situations. Basically, the program enables the investigator to examine the context in which a given word is being used within the text. Then, if he should find that the statute wherein the word is found needs to be looked at as a whole, the program provides a listing of the statute.

Therefore, once the investigator has selected the word or words (including misspellings) for examination, they are keypunched into IBM cards according to a specific format and read into the computer. The computer searches the vocabulary tapes for the occurrences of the words and notes their location (document, sentence, line, and word placement) on a scratch tape. This list is then compared with the text tapes and the line is printed out, with the selected word in the center and an equal number of characters (numbers, punctuation, letters) printed to the left and to the right. This provides the investigator with the word in the context in which the word is used. To the extreme right are found the document locations; volume, article, and section numbers.

As shown by several examples in Chapter IV, the KWIC program provides the investigator with a procedure, that if attempted by the normal hand search, would be virtually impossible to accomplish. For all practical purposes, an investigator would have to read every word of every page in order to obtain comparable results.

When the inquiry had been framed according to the required machine format, it wasthen punched on IBM cards. Using the IBM 407 printer, a printout was made of the inquiry and returned to the framer. Proofreading

of the printout is mandatory to check if any errors have occurred as a result of keypunching. Special attention was paid to original misspellings, since they must appear in the search list misspelled when fed into the computer. Any errors that were detected were noted, the appropriate cards removed and repunched, and were merged into proper, sequential order.

The inquiry was then merged with the control cards and "program deck," and submitted to the Computation and Data Processing Center for the scheduled run.

Professional Evaluation of Inquiries

In most instances, prior to the initial framing of the inquiry, an attorney was assigned to research the inquiry through the orthodox procedure of hand searching the appropriate sources (Figures 3 and 4). This searcher customarily worked independently of the framers. The rating system used in the analysis was devised by Mr. John Horty and his colleagues of the Health Law Center. By an arbitrary and human judgment, the documents found by hand searching are categorized as to their relevancy to the question. The citations found and printed by machine are also inspected and are similarly categorized as to their relevancy. Although quantitative analysis is in the early stages of study, the preliminary results have been relatively satisfactory, also. All the documents found were arbitrarily categorized, either by hand or machine, into three groups, "A," "B," and "C." "A" documents are ones that need to be found regardless of their application, for in this case, to an attorney <u>exclusion</u> is just as significant as <u>inclusion</u>. Or, knowing that a document exists that narrows the inquiry is as important as

For manual search #81 a partial listing of index entries would appear as:

CONSTITUTION — Education

Art. 10, §1. Duty of GA to provide public education
Kindergarten - Nothing
High School - Nothing
School Districts - Nothing applicable
Superintendent of Public Instruction - Art. 4, §20

INDEX — Kindergartens

24 §5-502 - Establishment
53 §16702 - Philadelphia
24 §25-2502 - Reimbursement from Commonwealth
Supplement 24 §15-1501 - Days of Instruction

INDEX — High Schools

Generally 24 §16-1601 to 1613
71 §352 - Classification
71 §354, 1023 - Course of Study - Duty to Determine
Directors, See School Directors
24 §5-502 - Establishment Joint High Schools, See
24 §16-1602 - New, Establishment
24 §1058 - Visitation
Supplement - Nothing Additional

INDEX — Board of School Directors

See School Directors
Supplement - Nothing

INDEX — School Directors

Repeat 24 §5-502
Admission of Beginners - 24 §13-1304
Assignment of Pupils - 24 §13-1310
Brands of Study Required - 24 §10-1005
Buildings, Duty to Provide - 24 §7-701
Closing Schools - 24 §13-1311, 24 §11-1153
County Board of School Directors
Sub-Chief Exec. Officers, Duties - 24 §9-926
Sub-Powers and Rites - 24 §9-925
Duties - 24 §5-501 to 5-523, 336 to 339
One Room Schools, Closing - 24 §13-1311
Reports - High School Information - 24 §16-1606
Special Schools - 24 §19-1921 to 19-1926
Supplement
Closing Schools - 24 §5-524

Figure 3

The body of manual search #81 would appear in part as:

#81 Pa. Constitution - Education

A. Art. 10, §1 - General Assembly shall provide for the maintenance and support of a thorough and efficient system of public schools wherein all the children of this Commonwealth above the age of six years may be educated. . . .

C. Art. 4, §20 - Superintendent of Public Instruction

Title 24 - Education

24 §2-211 to 2-217 - Powers & Duties of School Districts

C. 2-211 - General powers
C. 2-212 - Corporate Seal
C. 2-213 - Right to sue and be sued
C. 2-214 - Competency as witnesses
C. 2-215 - Conveyances by districts
Supplement 2-211 to 2-215 - Nothing
C. 2-216 - Gifts to Districts, Investment Accounts (and Supplement)
C. 2-217 - Unlawful gifts
Supplement - Nothing

24 §5-501 to 5-523 - Duties & Powers of Boards of School Directors

A. 5-501 - Elementary Schools - Board shall establish, etc., a sufficient number of elementary public schools... to educate every person residing in such district between the ages of six and twenty-one years.

A. 5-502 - Additional Schools and Departments - In addition to the elementary public schools, the board of school directors may establish, etc., the following additional schools or department ... High School ... Kindergartens.
Supplement - Nothing Additional

A. 5-503 - Kindergartens - (Supplement)-The board of school directors ... may establish and maintain kindergartens for children between the ages of four and six.

C. 5-504 - Cafeterias - (Supplement)
C. 5-505 - Establishment and Maintenance of Consolidated Schools
C. 5-506 - Duty to Consolidate Schools
C. 5-507 - General Powers, Taxation

Figure 4

knowing that a document exists that is directly relevant to the question. "B" documents are ones that could be considered of marginal value; they may be relevant, but again, they may not. The immediate visual inspection cannot determine this. Possibly, later events may make a "B" document important to the question at hand. "C" documents are those that visual inspection immediately shows are not relevant to the search.

Another semantic distinction must also be recognized. The documents themselves are arbitrarily divided into two groups, although the words employed usually have different meanings. "Sections" are treated as separate documents for the initial inspection. "Statutes" are treated as a composite of related sections on the basis that if access is made to one document, the others could be found. This division is necessary since a group of laws is usually passed by the legislature under a single act. Therefore, the arbitrary breakdowns are necessary in order that the searchers may work with reasonable categories.

This grouping of results into categories provides only one method of evaluation. In "practical" situations, the results are gauged by the simple question, "Do the results enable the lawyer to go ahead with appropriate action on a sound, legal basis?"

Another means of evaluation can be furnished by the non-attorney inquirer if the results provide him with an answer to his problem. For example, the senior author wished to know what statutory provisions were present to enable research in the broad field of education to be undertaken. The inspection of the machine printout, without the aid of an attorney, was sufficient to obtain an answer. The categories could be "relevant" and "irrelevant." Similarly, results of the KWIC searches

may be inspected by either an education specialist or an attorney to determine the success of the inquiry.

In summation, both the attorney and educator can and do contribute to the evaluation of the searches, with each respecting the competencies of the other. By this team effort, responsibility for evaluation is undertaken jointly in some areas, or independently in others.

## CHAPTER IV

## RESULTS AND DISCUSSION

In all twenty inquiries of value to educators, and others making decisions about schools, were run as machine searches. The titles of these searches and their standard inquiry search numbers are found in Figures 5 and 6. These inquiries involved the use of statutes and Attorney General's Opinion. The methods of restructuring the questions were described in the previous chapter. Legal opinions as the result of the conventional hand searches by attorneys were made for the majority of the inquiries. For the remainder of the inquiries an evaluation of the computer generated documents were categorized by one of the researchers (a non-lawyer) into two categories, important and unimportant. (In some instances both an attorney's search and one of the researchers made value classifications on the results of the search and the inquiry.) Figure 7 gives the statute and section counts by the attorney according to the "A," "B," "C," classifications (where applicable) and the machine search results. (The search numbers are nominal and are irrelevant for any other purpose.)

As the result of the close ties with the State Department of Public Instruction, the Attorney General's Office, and other state sources, nine inquiries received from these sources were run. The institutional training course on methods of framing inquiries also produced a number of questions by persons connected with the University, local school solicitors, and school administrators. Further, the training sessions in the school law classes produced additional inquiries. Some came from these first two sources as the result of general publicity in the state and from

LIST OF STANDARD SEARCHES

E 9 "May a school district invest surplus funds? If so, are there any restrictions?"

E 10 "Can a school district establish a junior college?"

E 12 "Is there any procedure whereby a school district may dispose (sale or otherwise) of unneeded or excess equipment or materials?"

E 14 "May two or more public school districts jointly purchase supplies and/or equipment, or construct and/or operate school plants?"

23 "What are the provisions for the health examination of school children?"

75 "What powers do the Superintendent of, the Department of, and the State Council have over private and parochial schools?"

81 "Is there anything in the law that requires a board of school directors to offer a kindergarten or to offer a high school program?"

83 "Can a school board buy liability insurance?"

89 "What powers does the state have in educational research and testing?"

179 "What are the powers and duties of the State Council of Education,
180 Superintendent of Public Instruction, and Department of Public
181 Instruction?"

183 "Is an Assistant Superintendent a public official under the law?"

196 "Is the guidance counselor a member of the administrative staff or is he a member of the teaching staff?"

Figure 5

## LIST OF KWIC SEARCHES

KWIC 1 "State Teachers' Colleges"

KWIC 2 "normal schools"

KWIC 3 "Supervising principal"

KWIC 4 "first Monday in July"

KWIC 5 "handicapped children"

Figure 6

## STANDARD SEARCH EVALUATIONS

| | SECTIONS | | | | | | STATUTES | | | | | | STATUTES RETRIEVED Only By | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Machine | | | Manual | | | Machine | | | Manual | | | Machine | | Manual | |
| | A | B | C | A | B | C | A | B | C | A | B | C | A | B | A | B |
| E 9* | | | | | | | | | | | | | | | | |
| E 10 | 0 | 0 | 5 | 0 | 0 | - | 0 | 0 | 5 | 0 | 0 | - | 0 | 0 | 0 | 0 |
| E 12* | | | | | | | | | | | | | | | | |
| E 14* | | | | | | | | | | | | | | | | |
| 23 | 15 | 2 | 26 | 15 | 1 | 1 | 7 | 2 | 22 | 4 | 1 | 0 | 4 | 2 | 1 | 1 |
| 75 | 32 | 2 | 25 | 20 | 4 | 13 | 19 | 2 | 20 | 15 | 2 | 9 | 4 | 1 | 0 | 1 |
| 81 | 7 | 30 | 62 | 6 | 10 | 8 | 6 | 15 | 41 | 5 | 4 | 1 | 1 | 11 | 0 | 0 |
| 83 | 5 | 2 | 68 | 5 | 0 | 27 | 5 | 2 | 53 | 5 | 0 | 27 | 0 | 2 | 0 | 0 |
| 89* | | | | | | | | | | | | | | | | |
| 179 | 87 | 2 | 0 | 51 | 0 | - | 65 | 2 | 0 | 38 | 0 | - | 27 | 2 | 0 | 0 |
| 180 | 243 | 6 | 14 | 148 | 0 | - | 93 | 1 | 8 | 65 | 0 | - | 30 | 1 | 2 | 0 |
| 181 | 214 | 12 | 21 | 153 | 0 | - | 105 | 8 | 7 | 66 | 0 | - | 55 | 8 | 6 | 0 |
| 182 | 39 | 2 | 38 | - | - | - | 16 | 2 | 13 | - | - | - | - | - | - | - |
| 183 | 0 | 0 | 23 | 0 | 0 | - | 0 | 0 | 23 | 0 | 0 | - | 0 | 0 | 0 | 0 |
| 196 | 0 | 0 | 41 | 0 | 0 | - | 0 | 0 | 30 | 0 | 0 | - | 0 | 0 | 0 | 0 |
| | | | | | | | OPINIONS | | | | | | | | | |
| 83** | | | | | | | 2 | - | 24 | 1 | - | 1 | 1 | 0 | 0 | 0 |
| 182** | | | | | | | 12 | 2 | 18 | 9 | 1 | 1 | 8 | 2 | 5 | 1 |

* See text for evaluation by education specialist

** Attorney General's Opinions

Figure 7

general legal research.

In general terms nine of the inquiries came from the authorities centered in the state capital: The Department of Public Instruction, the Attorney General, his Deputy for Education and his staff; ten from university centered persons, namely, university faculty and students of school law; and two from local school districts sources, school superintendents and school solicitors. These figures probably represent a reasonable approximation of the users of mechanical search procedures under general service conditions. It was found that the materials available, statutes and attorney general's opinions, but not judicial decisions, lend themselves well to the types of problems with which state authorities and students and teachers of school law are concerned. The local school district person is of necessity more oriented to answering specific questions. This is likely to require knowledge of court opinions as well as statutes.

Accuracy, completeness, and omission of irrelevant information is the test of the value of legal knowledge. Generally, tests of this type fall more into the purview of the research area of legal scholars and the originators of these procedures (6, 8). Those who are concerned with this aspect of the general problem are referred to the reports of these researchers. However, evidence needs to be presented to indicate that in this particular area of law, the results of this method of retrieval do not differ from those found in the general area of law. Again reference is made to Figure 7. Here is presented the various types of comparisons made in this study. As was indicated, this examination was made for eleven of the searches by the standard procedures employed in

the more comprehensive research project in law (6, 7, 8). This procedure requires the services of an attorney who, independent of the machine search procedures, examines the inquiry in the time honored manner of searching the law through the use of the standard indexes and guides. The categories of his searches versus those found by the machine searches were noted in the figure above.

A less sophisticated procedure was followed in evaluating the results of the machine search for the remainder of the inquiries. Here the investigators (neither of whom is a lawyer) rated the documents retrieved as relevant or irrelevant. The objective of this procedure was to evaluate the effectiveness of an unsophisticated person to learn to "frame" questions in the format required by the machine search procedures.[1]

The above results and referenced figures display in a rather mechanical way the effectiveness and efficiency of machine retrieval of documents in the field of school law. However, one needs to go beyond these figures to get to the total problem of the research. Just how feasible are these procedures for school law? This question can probably be best answered by examining the written evaluations, the case studies of each of the twenty inquiries that were used in machine searches during the course of this research. These will follow. For the convenience of the reader they will be identified by the search number and by the statement of the original inquiry.

---

[1] Since the expiration of the contract, the Health Law Center has prepared a Frequency List and directions for framing and submitting searches in less complicated format (9).

E 9 "May a school district invest surplus funds? If so, are there any restrictions?"

This search was a result of an inquiry submitted by Mr. Niles Anderson, Solicitor for the Board of Public Education of Pittsburgh. In this search, the question was not one of finding an unknown answer to a problem, but to compare a machine search with a hand search previously done by Mr. Anderson. The framer of the inquiry worked independently of Mr. Anderson, with the latter providing the final evaluation after the results had been analyzed. By allowing a non-school law expert to frame the inquiry and to evaluate the results, it was anticipated that some evidence could be made available as to whether a non-legal expert could properly frame and evaluate questions dealing with school law.

Instead of using the evaluative criteria of "A," "B," and "C," classifications described earlier, the education specialist merely categorized the results into groups of "yes" and "no"; yes, they are important and no, they are not. The conclusion was that although surplus funds could be invested and that there are several basic restrictions (as to the type of investment and the kind of institution), the problem was not fully resolved. This hinged on the matter of carrying a surplus from year to year, since technically, revenues are to cover yearly expenditures. However, for certain types of funds, as for future construction and purchase of sites, this would not apply. Such was the conclusion drawn by the education expert.

Similar conclusions were reached by Mr. Anderson, but his conclusions were more detailed. Furthermore, two important statutes were not obtained by the machine search, but were found by the solicitor. This

was due to the nature of the inquiry, which was framed too narrowly. In other words, the search was limited to the concept of school districts and their controlling agencies while the two omitted statutes referred to municipalities' sinking funds. But in reviewing the original machine inquiry, this omission was justified on the basis that had the inquiry been broadened, the expected output would have yielded far too much material in proportion to what might be elicited as important.

In this example, an important conclusion was drawn; an education expert could analyze a search in order to arrive at a tentative conclusion. But in an area such as the above, he must depend in the final analysis upon the legal advice of the school solicitor. Nevertheless, the approach is basically sound and provides a school expert with greater information than he could expect to find by means of conventional legal searching.

E 10 "Can a School District Establish a Junior College?"

The origin of this search can be traced to a request for inquiries from the university work group. In Pennsylvania the need for junior colleges is evident. Compared to some other states, the movement toward establishing junior colleges is in its infancy. One of the first requisites in establishing junior colleges is a knowledge of the legal authority, or lack of it, for an educational unit to establish this level of educational activity. It is a concern of local school administrators since many recognize the lack of a local higher education unit. Some of these men are represented in school law courses given at the University of Pittsburgh. Because it was a question of relevance and interest, it was a good test of the classes' ability to framequestions. It also

allowed the researchers to observe their strategic moves in doing so.

A condensed list of possible questions had been presented to the school law class and was narrowed to four or five questions. The above question was selected because it appreared to be the easiest for framing by the class and because the word frequency list showed a rather small occurrence of the words to be included in the concept of "junior colleges."

An analysis of the results quickly yielded the information that the law was silent in this area. Obviously, it is as important to know that the law is silent as it is to know it is explicit on a point. Therefore, the search was deemed valuable.

It provided an opportunity for the school law class to become participants in the project. They showed that even with a brief presentation of the methods of framing questions a representative group of present school and future school administrators learned the technique. Most also recognized the strategy of selecting words with which to frame questions from those which were relatively low in occurrence on the frequency word list developed in the original concordance procedures.

E 12 "Is there any procedure whereby a school district may dispose (sale or otherwise) of unneeded or excess equipment or materials?"

A request made by Mr. Lee S. Ellsworth, Associate Superintendent of the Franklin Area Joint School System, Murraysville, Pennsylvania, for answers to a list of questions was received through one of the non-investigator professors at the University of Pittsburgh who resides in the school district. Mr. Ellsworth had listed six questions with the above having top priority. The inquiry was framed by an education specialist, and a search completed on both the Pennsylvania Statutes and

Attorney General's Opinions. The evaluation revealed considerable documentation for the purchase of equipment and supplies, but none for their disposal. This information was made known to both Mr. Ellsworth and the solicitor of the school district, with the latter confirming this information as the result of his own searches.

The legal considerations drawn were that since no statutory provisions existed, the school district could either proceed on the assumption that because it was not forbidden to dispose of excess materials, the district could do so. Or, the district could not do so since no expressed authority was given to them by the legislature.

The following letter received from Mr. Ellsworth is evidence of the benefit of this particular search.

June 25, 1962

Dr. Theodore Polk, Director
Graduate Study
University of Pittsburgh
Pittsburgh, Pennsylvania

Dear Dr. Polk:

Thank you sincerely for answering my question on school law concerning the method of disposing of school equipment.

We had searched the law and found no mention of the answer requested.

One of our local state congressmen has promised to present a system for disposal of equipment to the legislature so that this problem will be legally answered.

(signed) Lee S. Ellsworth
Associate Superintendent

In this case, an important aspect of school law was found to be deficient and the computer search confirmed the search done by hand. Furthermore, the computer search covered all the statutory provisions, plus the Attorney General's Opinions which are not available unless

the searcher travels to the state capital.

This search confirmed the hypothesis that the machine method of inquiry can be used to the advantage of local school districts. Furthermore, communication of the limits of the machine searches do not necessarily cause local districts to be skeptical of its practical application. E 14 "May two or more public school districts jointly purchase supplies and/or equipment, or construct, and/or operate school plants?"

While investigating the feasibility of surveying a local school district, Dr. Richard DeRemer, a faculty member at the University of Pittsburgh, desired to have answered a question which could possibly arise in making school survey redommendations. Since the issue was not pressing and the explicit opinion of a legal expert was not needed, the question was used as a test of framing an inquiry.

After meeting with Dr. DeRemer, one of the investigators prepared an inquiry which was revised after further consultation. The revised inquiry was then submitted for a computer search in the morning. Later, while researching another school law problem, the education specialist assigned to frame the inquiry found the exact two statutory provisions which answered the question. Turning to the index, he then attempted to find these two provisions by the conventional means that a non-attorney would use. He failed. While a conventional search did not produce the answer, a pre-established check was available by which to compare the computer search.

The printout was later analyzed and evaluated, with the two provisions "on point" included in the computer generated material. An affirmative answer to both parts of the above question was found.

The importance of this search lies not only in the issue of having found an answer to a question, but also in that non-attorney education professors of school law interested in doing research in school law can be greatly aided by the relatively simple techniques developed for information retrieval.

While the question involved in this illustration was not one that required an immediate answer; nevertheless, the important conclusion is that research in school law can be done and with relative ease by professional people in education. This releases the professor from the time-consuming chore of hand searching and frees him for the more important task of analysis.

23 "What are the provisions for the health examination of school children?"

This search is of interest for two reasons; first, it is an example of an inquiry made by a non-investigator faculty member interested in possible research in this area; and second, a particular statute was retrieved that probably would never have been found by a conventional hand search.

Figure 7 notes the comparative figures between the machine and hand searches. There are two documents, one "A" and one "B", which were not retrieved by the machine. The "B" document refers to a statute dealing with Health Officers of Third Class Cities. Although nothing specifically relates to examining school children, the health officer is directed, as a duty, to give notice to the school or to the teachers whenever a child is stricken with a communicable disease. The health officer is usually made aware of the condition by a notice from the examining physician. This statute, then, implies the possibility of a health examination.

The "A" document refers to a section authorizing reimbursement to school districts for "School Health Services" but does not note that specific examinations are to be made. It states only that the Commonwealth will provide for reimbursement. Therefore, the "A" and "B" documents deal with general provisions that the investigator would want to be aware of, yet are not crucial, for specific documents answering the question were found.

The second reason noted earlier is of special interest. One statute was retrieved by means of the machine search from the Insurance Company Law of 1921, governing group health and accident insurance. Within the framework of this law is a provision allowing schools, or other institutions of learning, to be issued a group policy. Furthermore, and of significance for this search, is the granting of permission to the insurer, the company, to "examine the person of the insured" by an outside party of the company's choice. In essence, another instance of providing for the health examination of school children was discovered within a section of the law that in all probability would never be perused by a lawyer interested in school law.

This illustration demonstrates the information that can be obtained when a comprehensive body of law is available for a computer search. The additional tape library resources of the Health Law Center were of particular value for this search and further serves to demonstrate the unusual dividends gained by using the computer.

Many school districts in Pennsylvania do have insurance companies make available accident insurance to the students in this system. Yet in all probability very few, if any, administrators and boards of educa-

tion ever realized that they in doing so essentially give permission to the insurers to give a medical examination to their students.

75 "What powers do the Superintendent of, the Department of, and the State Council of Education have over private and parochial schools?"

A professor became interested in this specific question as a result of having been involved in another closely related area of investigation. He made contact with staff members of the Health Law Center for assistance in framing this inquiry. The initial attempt at framing was completed but was not punched onto cards for machine inquiry. It had become evident to a member of the Center that the inquiry had been framed too narrowly and that it would have to be expanded to cover a broader concept of the word "private." In effect, the word had to include not only private schools dealing with a school age population, but various commercial educational enterprises for the adult population and which required licensing by the state.

This search was the initial encounter by the investigators with the framing techniques that blended and united various concepts according to instructions fed simultaneously into the computer with the key words and their combinations. Prior to this search, the investigators had been familiar only with the use of the "or" indicator (a "+") noted in Chapter III.

This search produced superior results. It delineated specifically, and in a "handbook" form, all the rules and regulations relating to private and parochial schools by areas of authority; the Superintendent, the Department, and the State Council of Education.

81 "Is there anything in the law that requires a board of school directors to offer a kindergarten or to offer a high school program?"

The above was one in a series of questions submitted by Dr. Neal V. Musmanno, Deputy Superintendent of Public Instruction, representing the type of question in which he had an interest. Since at the time the investigators had not yet become proficient in framing and/or analyzing inquiries, an attorney was first assigned to do a manual search of the question. At a later date, the junior writer was assigned to prepare the machine inquiry to become familiar with the procedures involved in legal information retrieval.

Working closely with a staff member of the Health Law Center who had not been involved with the manual search, the question was discussed and revised. Later, the question, now in machine form, was keypunched into punch cards and fed into the card reader along with other inquiries. The printout was available approximately one hour later. Meeting with the attorney who had prepared the hand search, the framer along with the second attorney, met to analyze the results.

Inspection and comparison noted a larger number of "C" elicited sections by the computer search than was the case with the manual search; 62 to 8. (See Figure 7) Such a discrepancy was questioned by the framer. The obvious answer was that the computer searches list all documents meeting the instructions of the framer while the hand seacher, consciously or unconsciously, does not investigate or note all documents which do not appear to be relevant. Further more, the computer search involved all the Statutes of Pennsylvania while the hand searcher would not be led into many sections of the code through the indexes. Finally, it should

be noted that any framer of an inquiry will willingly accept a larger number of "C" documents providing the "A" and "B" ones have all been elicited. This excess material can easily be inspected and quickly discarded if found to be of no assistance to the question under investigation.

In summation, this question not only provided an opportunity for the junior writer to become actively involved in framing and analyzing, but it also reflects the position that a computer generated search, while eliciting a larger number of irrelevant documents, is far more successful due to its thoroughness and availability for immediate visual inspection.

83 "Can a school board buy liability insurance?"

A search on the Attorney General's Opinions was run in hopes of further clarification of the statutes. While only two "A" opinions were found by machine, only one was produced by the conventional search procedure.

This example illustrates that even a yield of one more pertinent opinion is of value due to the length of time usually required to hand search the opinions. Statutes, being relatively brief, can be perused rather quickly, but opinions from the bench or the Attorney General contain much verbage which must be closely inspected for any clues relevant to the legal point.

In essence, then, the number of documents or opinions retrieved is often overshadowed by the simple fact that they can be made accessible through a machine search.

89 "What powers does the state have in educational research and testing?"

The search was made as the result of the senior writer's interest in supporting and promulgating research and research support services of the state. Knowing the value which governmental fiscal and budget personnel hold for statutes and Attorney General's Opinions, it was felt that every statute and interpretation of statute, that even remotely authorized or directed the state to perform a research or research related function would be of value in persuading the controllers of the purse strings to allow a more active assistance in this vital field.

The inquiry was framed very broadly by the junior author to cover every conceivable idea relating to these functions. It was known that the output would be relatively large, yet it was felt that the material would be difficult to locate in the usual search procedures and that the output could be read rapidly and discarded if found irrelevant.

The output was indeed extensive; it was one of the highest cost questions put to the computer during the project and consumed 139 minutes of time at a cost of $347.00. The output on 8½" by 14" paper was about 2½ inches thick.

Using the classification system "A," "B," and "C" to describe degrees of relevancy of the retrieved documents, of the 104 documents retrieved 101 were irrelevant. These ranged anywhere from documents indicating that tests could be given to physicians desiring to become licensed to practice in Pennsylvania to forms needed for filing reports with state agencies. Two "B" documents were found and just one "A." However, this was an extremely valuable find for it was an Attorney General's Opinion which construed the equalization of educational

opportunity statute to mean that funds could be expended for education research since the results tended to indeed improve educational systems generally. These in turn, were interpreted to mean that they created a more equal opportunity for school children to receive a good education.

This opinion was a particularly important discovery because it is practically impossible to find by conventional means. It is not indexed in any of the usual information retrieval systems commonly used by attorneys.

The two "B" documents were resolutions of the legislature commending educational research. While these resolutions do not have the force of law, they are valuable in attempting to persuade future legislators that educational research has been looked upon favorably in the past. Fiscal and budget personnel in the state government also tend to look more optimistically on legislative and budget requests if they believe that the legislature will receive them favorably.

This particular search is evidence of the value of varying breadths of searches for various types of consumers. The legal investigators doing research in the general area of legal information retrieval might will be aghast at an inquiry which yields 101 irrelevant documents to one relevant one. However, to a non-legal scholar working in the field of law the mass of documents was not of major concern. Any other method of inquiry would have been far more costly.

The investigator in his role as consumer deems this result as worthwhile and valuable.

179 - 180 - 181 "What are the powers and duties of the State Council of Education, the Superintendent of Public Instruction, and the Department of Public Instruction?"

These questions were submitted by the Deputy Attorney General to serve several purposes. In Pennsylvania there are three administrative bodies which have authority over education at the state level. These are the Superintendent of Public Instruction, the State Department of Public Instruction, and the State Council of Education.

It was of considerable legal interest to know where there were overlapping, co-existing functions, and specific functions of each body. Finally, it was felt that a "handbook" of the printout for the statutes dealing with each body would be a valuable aid to the Deputy Attorney General who has major responsibilities in answering questions arising out of the interpretation of school laws.

There were three immediate purposes. First, the intent was to obtain a total compilation of all instances wherein any of the three agencies were involved, as mandated by the legislature; second, such a compilation would provide an immediate reference source for specific questions submitted to the attorney general's office for advice. Finally, the enactment of a new school reorganization bill could involve the drafting of new legislation, redefining certain powers and duties of the three controlling agencies and the knowledge of present legislation could be crucial.

The general evaluation of these searches proved that documents were retrieved in sections of the law that any diligent hand search could not have uncovered, unless a page-by-page review was done of all 75 volumes

of Purdon's Pennsylvania Statutes Annotated. For instance, even though the Deputy Attorney General had at the time considerable experience in the position, he was unaware -- as was probably the Superintendent of Public Instruction -- that the Superintendent has certain obligations under the narcotic and cocoa leaves sections of the Pennsylvania laws. The computer did not miss these sections however. This question once again demonstrated the need to have available for machine search all the statutes and not just the usual Education Code. Furthermore, Mr. Killian the Deputy Attorney General reports that he feels quite confident that the results represent as complete a listing as could have been obtained under any circumstances.

As a result of these searches, one needed change in the reporting of retrieved information was made obvious. At the time due to costs involved in printing, the material sent to the Attorney General's Office contained only the citations to the statutes. This was necessary because of the increased cost in printing out the entire text.

With the installation of a separate computing complex, devoted solely to information retrieval, and the installation of a more rapid printer operating on a separate line, complete text printouts are now customary. Therefore, in a matter of minutes, the printout may be examined without the time consuming procedure of "paging" in an annotated volume of statutes.

182 "What are the provisions for leaves of absence for school employees in the Attorney General's Opinions?"

In order to obtain a more complete picture of provisions governing leaves of absence for school employees, a manual and a machine search

were done on the Attorney General's Opinions. It was desired in order to clarify certain statutory provisions.

The comparison of the machine results with a hand search (Figure 7) revealed that the manual search yielded six opinions, one "B" and five "A," not found during the machine search. This descrepancy necessitated further examination with the following results. The "B" opinion referred to "Civil Service Employees," which does impinge on the area of teaching employees yet does not directly relate to them. The five "A" opinions revealed that two dealt with "State Employees," one with "Regular Employees," one with "Women Employees," and the last with "Rights of returning veterans."

Even a superficial reading of these headings revealed that all were broader in scope than was applicable for this problem. In other words, the opinions not found by means of the machine search were of a general nature, having applicability only if specific opinions had not been available. Generally, it is accepted that a specific ruling over-rides a general one. Therefore, one could conclude that while the six opinions retrieved in the usual searching procedure were ones of which we would wish to know, their loss would not necessarily negate the conclusions reached.

The results of this search again demonstrated that the machine search of the Attorney General's Opinions is certainly easier than any form of manual search since none of the usual indexes are available for hand searches. Also, the breadth of a search by the framer may be too narrow to obtain relevant opinions. The latter is always a hazard of the non-lawyer making decisions in areas based upon law. The same precautions

that apply to the professional educator's interpretation of the law also apply to his use of machine search methods. As is typical of almost any use made of computers, when the intellectual effort directing their use is explicit and excellent, the results are explicit and excellent.

183 "Is an Assistant Superintendent a public official under the law?"

This problem came about as a result of the desire of a county district to raise the salary of its Assistant Superintendent. However, the state constitution specifically prohibits raising the salaries of public officials during their term in office. The question was then posed as to whether the assistant superintendency was an office falling under this prohibition. A search of the statutes had yielded 23 provisions within this general area, but none answered the point in question.

Turning then to the Attorney General's Opinions, a machine search generated four opinions: two of which specifically dealt with the problem and two supported the preceeding decisions. A manual search was undertaken and only one of the specific opinions was found in this manner. An enigma arose, however. The provision of the state constitution which brought up the questions in the first place, was not among those retrieved by the machine search.

Obviously, in an area where in the statutes are silent and an interpretation is crucial, it is not only desirable, but necessary to have a swift and reliable procedure for ascertaining whether such information is in existence. The results of the search were of immediate value to the Deputy Attorney General. To the investigators, it was one more piece of evidence concerning the value of machine search procedures in school law.

196 "Is the guidance counselor a member of the administrative staff or is he a member of the teaching staff?"

The framing of this inquiry was an outgrowth of discussions held in the school law class. Several of the students working in a study group, decided to seek an answer to this question since they felt it represented an easy question to frame. This was due to the very low frequency of the words "counselor" or "counselors" (14) and would present them with an opportunity to examine a computer-generated output of their own selection.

They presented the investigators with the framed inquiry, the search was done manually and by machine, and the results returned to the class for discussion. As shown by Figure 7 no "A" or "B" documents were retrieved by either method. The conclusion is that the statutes were silent on this point.

Unfortunately, at this early stage of the investigation, the Attorney General's Opinions were not yet available on magnetic tape and a machine search of them was not done. The students were cognizant that in such circumstances, wherein the statutes are silent, the only possible resolution in determining an answer rests with case opinions and Attorney General's Opinions.

As a teaching device for school law courses, it is felt that computer methods are reliable. This problem was submitted in the early weeks of the school law course. The students received valuable insight into the relationship of the lawmakers and the interpreters of the law. Further, it was obvious that students quickly became aware of the strategy as well as the mechanics of the method of framing questions.

KWIC 1 "State Teachers' College"

On May 25, 1961, P.L. 227 was passed by the Pennsylvania State Legislature. Part of this Law changed the official title of the several State Teachers' Colleges to State Colleges with prefix words identifying the city in which the college was located; e.g. Slippery Rock State College. Each occurrence in the statute needed to be changed. Any oversight of a statute or missing of one citation in a statute while correcting another could cause legal confusion. All the corrections had supposedly been made.

In this example, several instances were found where the repealed word "Teachers" appeared, despite the careful perusal by the legislators, their legislative assistants, and the legal counsel of the Department of Public Instruction.

This is an example of the assistance of the KWIC program might be in framing legislation. A search was run on the word "college." A quick visual inspection of the words immediately preceeding "college" would then have indicated the provisions needing amendment.

KWIC 2 "normal schools"

If even "State Teachers' Colleges" no longer has legal meaning, why are there still several references to "normal schools" remaining in the statutes? This search was an outgrowth of the KWIC 1 search which was to determine the occurrences of the phrase "State Teachers' Colleges." At the time the phrase was to be searched by the computer, a review of the index of Title 24, Purdon's Pennsylvania Statutes Annotated revealed that there were still several entries under "Normal Schools."

The question was then raised as to why this entry was still included since the 1949 revision of the School Code was to have eliminated this phrase and substituted "State Teachers' College." This search was done to determine the remaining statutory usage of the phrase.

Analysis of the search revealed that twenty-one occurrences of "normal" referred to Normal Schools and Normal Districts, with most of the remainder pertaining to certification and/or diplomas. The latter is to be expected because teachers who had attended Normal Schools would have received such certification. However, a few other occurrences were errors of oversight in the amending procedures. This is but another example of the routine procedure of hand searching resulting in possible legal ambiguity.

KWIC 3 "supervising principal"

KWIC 1 and KWIC 2 were run after the fact. The legislature had acted, and the searches were intended to locate errors of omission in the enactments. In this example the search was run, it is hoped, before the fact. In Pennsylvania, the duties of a supervising principal are, almost in fact, that of a superintendent. The multiplicity of terms is confusing. It has been desired for some time to consolidate the two. However, there are a considerable number of statutory references to the term "supervising principal." To amend the considerable legislation on the record would demand a great deal of legal preparation on the part of the Department of Public Instruction. Thus, an improvement in the administrative structure of schools' systems is held back by legal insufficiencies. Actually, a much needed school district reorganization in Pennsylvania is somewhat dependent upon these changes.

Occurrences of the phrase "supervising principal" were sought prior to any considerations or deliberations by the legislature. It was thought that, if the legislature would desire to change the title (and certification requirements) of supervising principals to superintendents of schools, such a change could be completely, accurately and immediately done by providing them with the listing of the occurrences. With a complete and accurate listing a minimum of legal confusion would follow the amendments.

The listing was made on one machine search -- along with several others -- and given to the Deputy Attorney General who is the legal consultant to the Department of Public Instruction. His comments indicated that the search was of extreme value. The pro rated cost of the run, assuming a tape library of documents, was on the order of twenty dollars.

KWIC 4 "first Monday in July"

At the request of the Office of the Attorney General of Pennsylvania, a search was attempted to find all statutes in which the phrase "first Monday in July" appeared. This information was needed since all governmental bodies in Pennsylvania except school districts begin their fiscal year on July 1. School districts begin their fiscal year on the first Monday in July. This causes unnecessary confusion, non-conformity, and difficulties in analyzing revenues and expenditures. In order to amend all the statutes to conform with the other governmental bodies, all "first Monday in July" phrases had to be identified.

According to the Attorney General's Office, an accurate listing of all sections to be amended were made available. With the context printout, an attorney could quickly mark the copy appropriately and give it to a bill drafter for implementation. While many of the legal searches

done during this research could have been done in the standard way, some perhaps even at less cost, this is one search which could have been done only by a complete word by word reading of the statutes. No conventional means of indexing is designed to locate this phrase. Even with the reading of the statutes, it is well known that no human mind can possibly perform such a long, dull task with any degree of accuracy. Much of the development of partial sampling procedures in industry inspection jobs was done to increase the precision of inspection. It was observed that it was better for people to inspect a small fraction of a group of products very closely than to attempt to inspect them all. The analogy holds true in the field of law. This simple, yet dramatic example illustrates the thoroughness and low cost which is possible in a machine search.

KWIC 5 "handicapped children"

The results of the research for this search are similar to those given in KWIC 4. However, two methods of machine searching were used, the standard and the KWIC search. Interestingly, the KWIC search, the one simplest to learn to program, was actually somewhat superior to a conventional search by a semi-novice.

As in the case of KWIC 4, the legislature wished to change all occurrences of "handicapped children" to the more accepted term "exceptional children." Two types of searches were done for this problem. The first produced document titles. The second search used the KWIC which provides the document title, plus several words appearing on either side of the word "handicapped."

The KWIC search was the most helpful, since not all statutes containing the word "handicapped" referred to children. Therefore, visual inspection of the KWIC search results immediately gave an indication as to the need to amend the statute.

## CHAPTER V

## POTENTIALS OF COMPUTER ANALYSIS OF DOCUMENTS

It is the belief of the writers of this report that the value of the computer analysis of school law documents involving legal materials has been demonstrated beyond doubt. In working with these materials during the course of the research, a number of other possible applications and uses of the computer for document analysis were conceptualized. Two of these uses of the computer would be for the codification of school law and the preparation of smaller codes within the larger field of school law.

The flexibility with which the documents and statutes within school law can be handled makes the computer methodology an excellent one to use in restructuring and reclassifying the existing laws. One example cited was that of the reorganization of school districts. In Pennsylvania the school law relating to reorganization of school districts is now found throughout the whole volume of statutes relating to school law. It is quite feasible to put all the statutes relating to a given general topic into one place first by searching for the general concept through the use of the computer and then putting the printouts of this one topic together in the form of a dummy school code. A number of states have an urgent need to recodify their school laws, and perhaps this is one of the easiest ways for the reorganization to take place.

The preparation of smaller, sub-codes is also entirely feasible through the use of computer analyses. For instance, it would be quite possible to obtain printouts of all the school laws having to do with

attendance, or with school buildings, or with school finance, or even such narrow areas as in the Pennsylvania laws as school law pertaining only to Class One Cities. A number of these smaller codes could be of value to, say, teacher workshops, attendance officers, and superintendents interested in building programs. In the case of the Class One City code, there are probably only a half-dozen men in Pennsylvania who are primarily interested in them. They are the men who interpret the school law for the two Class One Cities in the state of Pennsylvania, Philadelphia and Pittsburgh. While in this latter case, these men are already quite familiar with this law, it might be valuable to them to have a printout of all the school law dealing with the Class One Cities especially when the relatively modest cost of obtaining it is considered.

It might also be valuable to the legislature to have an indication of the areas in which the law is silent in certain classes of school districts.

As the computer analysis of school law (or general law for that matter) becomes established, the attachment of a Flexowriter or other device to a print setting machine would simultaneously punch cards or mark a tape which could then be converted into the magnetic tape code of the computer. As new laws are passed and are printed for the public record, it would be possible to make copies of the laws for insertion into the tape library for computer use. This would reduce the cost of the input procedures and also make immediately available all of the school code of a given state.

Another possible use would be the development of new laws pertaining to schools. It might be possible to have in the magnetic tape library

the laws from several states. Concepts from one state which might be applicable to the educational process in another state could then be used, produced, and duplicated rapidly for legislative guidance. The best law or the most valuable combination of sections could then be enacted.

To turn away from the field of law for a moment, it might be interesting to speculate about the use of document analysis in fields such as science, communication, and the field of education, per se. Indexing and abstracting are an inherent part of the communication process in the field of science. It has always been a trying and difficult area, one in which scientists are concerned because of the time lags involved between the publication of a report and its abstracting and indexing for reference purposes. The lag between the publication of an article and its abstract can run as much as 2 to 3 years during which time it is not generally available to other scientists. In the meantime, of course, it is only by gratuitous circumstances that a scientist becomes aware of the information which is contained in the published article. The abstracting of an article generally is of such nature that it must be done by professionals in the field. This is a costly process, and it diverts scientists from their major effort of obtaining new knowledge. To have someone other than a professional attempt to abstract an article often means that the abstract loses some of the major concepts of the journal article. Even in attempting to find key words which will adequately retrieve an article, it is a costly process to which no good solution has yet evolved. It has been the experience of the writers in doing the research for this publication that it may be cheaper to

put the entire article into a magnetic tape library, or at least take an abstract prepared by the original author, even though it may seemingly be inadequate for the very precise information retrieval purposes which are called for by present day techniques. It might even be possible to take a general introduction to a scientific article including a statement of the problem and the summary and conclusions of the article since these sections can be quickly determined by professional persons at a very low cost. The actual translation from printed page to magnetic tape library can be done relatively cheaply by semi-skilled labor. If by any chance the summary and conclusions or the statement of the problem does not include major words which would be of interest to other researchers, it is no problem for the professional to quickly add these key words individually at the end of the journal article, if necessary, out of the context of the article. They need not be in sentence or paragraph form. This would be particularly important if the abstract of the article is prepared, as it is in many scientific fields by the original author. The original authors may not put into their abstract the ideas or concepts necessary for efficient indexing or retrieval. They may tend to omit important words or phrases from their abstracts. These important words or phrases could be quickly added at the end of the abstract even though they would seem not to make sense to the ordinary reader. However, the original abstract plus these added words and phrases would be invaluable for retrieval purposes.

At the end of a year using the abstracts and these key words and phrases from the abstracts, it would be possible to prepare a concordance of all the words used in the abstracts. This would be complete

and ould give the exact reference location of each one of these words on a particular page of the journal. None would be lost and the complete year's index would be immediately available at the end of the year. Series of years; one, two, three, five, ten, would be merged together at the end of each span of years so that a block of years could be indexed. Instead of going through ten separate indexes, year by year, a ten year's block of index words would be available in one place.

It would also be possible by putting the titles, the authors, and the abstracts on different kinds of cards, say type one, two, and three cards as used in this current system, to make a complete alphabetized index of author's last names and article titles. Again this would give the complete references to the output for the year and over a period of years.

While this procedure would not be particularly more economical than the present procedures for any one year, the real advantage comes over a period of years when two, three, five, ten, twenty, and more year indexes can be made with negligible effort and which are known to be complete. Thus it would be possible to get the complete writings of a given scientist during a ten year period simply by finding his name in the ten year index.

It would also be possible to obtain all the articles written about a given concept such as listening, delinquency, or schizophrenia. Since many of these concepts are quite broad and would yield a large number of titles, the reference location or abstracts, as the researcher wished, could be elicited. It would be quite possible to delimit some of these concepts by making a combination of say schizophrenia in childhood by

calling for schizophrenia plus adolesence, plus infancy, plus any other synonym for people less than 18 years old. It would thus be possible to do rather complete literature searches mechanically and make them available to researchers.

Another use which has been considered for the analysis of documents has been vocabulary lists. The Thorndike word lists are probably most commonly used in the fields of education and psychology today although there are others. The Thorndike lists are almost 30 years old, and they are counts of printed words only. There is no particular classification of these words. This should not be viewed as a criticism of this magnificient study, since it is extremely valuable, but as an indication of how costly the original study was. The study was done during depression days with WPA labor. It probably has not been replicated because of the expense. The computer could be used to make new vocabulary lists which would be more up to date. These vocabulary lists could also be made from various types of documents such as newspapers versus magazines. They could be used for determining the word list of various types of reading which is done by various types of groups such as first graders, graduate students, delinquents, or upper socio-economic groups.

It would also be valuable and entirely feasible using computer techniques to make an analysis of the <u>speaking</u> vocabularies of various types of groups, such as the deaf, the mentally retarded, or a particular age or grade level. Already the senior writer of this report has started along with a colleague to work on the vocabulary list of first grade books.

To these writers the most obvious limitation of the computer technique particularly in the field of law is that it will not make human decisions or interpretations about the law. This is obvious from the reports and the results of the studies and searches that were made. Thus lawyers need not feel "threatened" by automation, but rather they should feel that the magnetic tape law libraries in the computer should assist them in doing an ever better job of practicing law.

# CHAPTER VI

## CONCLUSIONS

A number of conclusions can be drawn from the results of this study. Generally, they fall into several categories, namely: Costs, Framing, Analyses, Users, Potential of computer use in school law, Potential of computer use in other document analysis fields, and the General conclusions.

### Costs

1. The preparation of materials to be placed on tape is a major cost of this system of information retrieval.

2. The initial investment of computing equipment for such an undertaking would be prohibitive to be replicated in many institutions.

3. The machine time needed to concord and provide frequency word listings for the form needed in searching is a matter of many hours of computer time.

4. The cost of searching is relatively inexpensive (compared to other methods) once the initial expenditures of equipment and concording is completed.

5. Changes of words in statutes by amendment can be done quickly, thoroughly, and inexpensively using the computer as a search device.

### Framing

1. The education expert can effectively develop machine searches and analyze the results within his own sphere of competency.

2. The process of framing inquiries can be taught to competent personnel within a relatively short period of time.

3. The word frequency list must be used in framing inquiries.

4. Professors interested in school law research can be aided by releasing them from hours of turning pages and scanning indexes.

5. Complete tape libraries are necessary to do an accurate and thorough machine search.

Analyses

1. A machine search almost invariably elicits the one or two more documents crucial to the question under investigation.

2. Even one opinion of the Attorney General retrieved by machine is invaluable, since many researchers would be unable to locate it by any other method.

Users

1. Those concerned with writing, amending, and enacting laws are the major beneficiaries of this research at this point of development.

2. The most frequent users of the tape library were state officials and faculty members.

3. Practitioners of school law, generally, will not be satisfied by citations only, full text printouts are immediately used while citations create delay since they must be located in book form.

4. Graduate students gain more insight into the inter-relationship of legal agencies by becoming actively involved in searches.

5. Those interested in specialized, minute details of school law can now avail themselves of smaller school codes generated by computer searches.

6. Generally, users want full text printouts to permit immediate application of retrieved information.

7. Education experts can, in many cases, only reach tentative conclusions and must await the judgment of the legal expert who investigates rulings and court decisions bearing upon the issue.

Potential of Computer Use in School Law

1. The attachment of a Flexowriter to a type setting machine could provide fast, inexpensive input for a tape library.

2. Materials, such as Attorney General's Opinions, which are not available to users at central repositories, are assets to any tape library. Using a tape library would be a valuable means of disseminating legal materials.

3. Legislators could avail themselves of concepts, wording, or phrases from other states in order to obtain the best phraseology for new laws when other state tape legal libraries become available.

4. Codification of school law using computer facilities is feasible.

5. The preparation of smaller school codes within the larger field of school law is feasible.

6. New programs must be written to take even more advantage of having a complete text library.

7. Undoubtedly even more efficient machine search procedures will evolve as a result of higher speed computers and as the result of the findings of the basic researchers in the field of legal information retrieval.

Potential of Computer Use in Other Document Analysis Fields

1. Indexing and abstracting of articles, speeches, or papers for use in computer searching would be a logical development in view of the results obtained in this project.

2. Multi-year indexes could be kept on tape and searched by computers according to a wide variety of categories; e.g., author, year, general title, key words.

3. Total concepts could be retrieved via an abstracted tape library which cuts across authors, years, or specific titles; e.g., delinquency, maturation, heredity.

4. Concordances of many types of word lists (written or spoken) are, for the first time, within the economic range of research, both in time and money; e.g., retarded children, early admission college freshmen, 7th grade history books.

General Conclusions

1. The human factors of intelligence, critical thinking, decision making, etc. are crucial and basic despite the complexity and degree of sophistication of the computing equipment involved.

2. The time lag between the completion of this research and its application in the educational process was small. This is in contrast to the proverbial 50 year lag.

3. The value of computer analysis of school law documents has been clearly demonstrated.

## REFERENCES


1. Brittell, Robert W., "The Use of the High Speed Electronic Computer as a Research and Operational Device in School Law," Speech given before National Organization on Legal Problems in Education, Indianapolis, Indiana. November, 1961.

2. Brownson, Helen L., "Research on Handling Scientific Information," Science, Vol. 132, No. 3444, December 30, 1960. pp. 1922-1931.

3. Campbell, John D. and Carson, Herbert S., "Data Processing by Optical Coincidence," Science, Vol. 133, No. 3461, April 28, 1961. pp. 1333-1338.

4. Horty, John F., "Application of Information Retrieval Techniques to Legal Research," Unpublished paper, November 1, 1960, University of Pittsburgh, Pittsburgh, Pennsylvania.

5. Horty, John F. and Kehl, William B., "Legal Information Retrieval at the University of Pittsburgh," Unpublished paper, University of Pittsburgh, Pittsburgh, Pennsylvania.

6. Horty, John F., "Electronic Data Retrieval of Law," Current Business Studies, Vol. 36, 1961. pp. 35-46.

7. Horty, John F., "The Key Words in Combination Approach," M.U.L.L., March, 1962.

8. Horty, John F., "Experience With the Application of Electronic Data Processing Systems in General Law," M.U.L.L., December, 1960. pp. 158-168.

9. Horty, John F., "Word Frequency List of the Statutes of Pennsylvania" (With Instructions for Requesting Computer Searches), Health Law Center, University of Pittsburgh, Pittsburgh, Pennsylvania.

10. "IBM General Information Manual; Introduction to IBM Processing Systems," International Business Machines Corporation, Data Processing Division, White Plains, New York, 1960.

11. "IBM Reference Manual; IBM 1410 Data Processing System," International Business Machines Corporation, Poughkeepsie, New York, 1961.

12. IBM Stockholder's Quarterly Report, July 11, 1962.

13. Kehl, William B.; Horty, John F.; Bacon, Charles R. F.; and Mitchell David S.; "An Information Retrieval Language for Legal Studies," Communications of the ACM, Vol. IX, No. 9, September, 1961.

14. Fels, Eberhard M., "Evaluation of the Performance of an Information-Retrieval System by Modified Mooers Plan," American Documentation, Vol. 14, No. 1, January, 1963. pp. 28-34.